Sabri A. Mahmoud,
"Motion analysis of multiple moving objects using the Discrete Sine Area Transform (DSAT)."
The Journal of King Saud University - Computer and Information Sciences Division, Volume 7, 1995, pp.25-48.

Key words: Discrete Sine Area Transform, Motion Analysis, Motion Estimation, Multiple Objects Analysis

J.King Saud Univ., Vol.7, *Comp. & Info. Sci.*, pp.25 - 48 (A.H 1415/1995).

Motion Analysis of Multiple Moving Objects Using the Discrete Sine Area Transform (DSAT)

Sabri A. Mahmoud

Computer Engineering Department, College of Computer & Information Sciences, King Saud University, P.O.Box 51178 Riyadh 11543, Saudi Arabia

(Received 13 June 1993; accepted for publication 12 January 1994)

Abstract. A new technique for velocity estimation and motion analysis of moving objects is presented. This technique is based on applying the Discrete Sine Area Transform (DSAT) to image sequences. The formulation of applying DSAT to an image sequence of multiple objects is presented. The analysis of the DSAT spectral domain indicates that the velocities of the moving objects are related to the locations of the peaks in the SAT spectrum. A peak detection algorithm is used to locate the DSAT spectral peaks from which the velocities of the moving objects are estimated.

Experimental results are included to demonstrate the applicability of this technique. This technique is simple, efficient, and gives accurate estimation of the velocities of moving objects. There are no limitations to the size nor to the velocities of the moving objects. The presence of stationary objects does not affect the presented analysis nor does it require special processing.

Introduction

The estimation of the velocities of moving objects is needed in many applications, such as angiography studies of the heart, biomedical cell motion analysis, tracking dust storms and clouds, and in remote sensing and military applications. The temporal- spatial gradient technique [1-4], which utilizes low level estimation of frames has been employed. These methods give best results only when the moving objects have smooth edges and the surfaces contain prominent texture. Their algorithm requires registered image sequences and the velocity estimate is approximate(it is close to accurate velocity at objects' boundaries, while this is not the case at the objects' surfaces). Such methods are subject to significant problems due to occlusion and the presence of certain surface textures.

Other researchers have used differencing techniques to extract moving objects in a sequence [5-7]. In their technique two frames are subtracted and the resulting frame gives the change between the two frames. This has limitations as it requires the images to be exactly registered, illumination to be invariant,and the moving objects to be totally displaced. Segment and match technique have been used to acquire velocity information [8-13]. Static images are segmented and then feature points are matched to establish correspondence between objects in successive frames. This technique is sensitive to segmentation errors and the success of the algorithm is based on accurate segmentation of static frames which is rarely satisfied in real world scenes; especially in noisy environment.

Optical flow, which is determined by obtaining the velocity vector for each pixel in the image, was used in [14-23]. A distinctly different combination of object structures and motion can produce effectively equivalent optical flow fields. This makes it impossible to un-ambiguously determine motion and structure from changes in the imagery [17]. Hence, flow fields estimated from changes in an image sequence are unlikely to be accurate. Jacobson and Wechsler[14] applied Wigner distribution in conjunction with a velocity polling function. Heeger[20] used a family of motion sensitive Gabor filters(3-d space-time Gabor filters tuned to different spatio-temporal frequency bands) to determine the displacement field. A least square method is applied to compute the two-dimensional displacement vector from the set of filter responses. Young & Kingsbury[23] used Complex Lapped Transform for frequency domain motion estimation where overlapping blocks of data are used to estimate local motion vectors. These data blocks, at each shift position, are first windowed by half-Cosine function. Their technique is free from block edge discontinuities which are present in other block matching techniques[24,25]. However, the technique is less straightforward to implement than block matching and thus have greater overheads. These techniques require high processing efforts and 3-d analysis is performed. The proposed technique apply similar analysis to [14,20] and hence may be used to replace the Wigner distribution in [14] and the Gabor filters in [20] with savings in processing time and memory requirements. The above techniques have other limitations when multiple moving objects are in the sequence; the correspondence and occlusion problems are more severe.

This paper presents a new approach based on the analysis of the Sine Area Transform spectrum of the image sequence for motion estimation. The Sine Transform [26-30] has found applications in image processing and is computed from the Fast Fourier Transform. It can also be computed from the Hartley transform as proposed by Hou [31]. In this technique, the DSAT is applied to the image sequence followed by a peak detection procedure. This technique may be used to obtain the optical flow of image sequences using the sectioning algorithms of[14&32].

The organization of the paper is as follows. The analytical model and formulations for multiple moving objects in a time-sequence are presented in section II. Section III presents the experimental results. Finally,concluding remarks are given in section IV. A detailed formulation of the application of DSAT to an image sequence with multiple moving objects is given in appendix.

Analytical Formulation

A general two-dimensional image sequence with time as the third dimension is given by [33],[34,pp.1-18] :

$$f(x,y,t) = O_m(x,y,t)\, o(x,y,t) + [1 - O_m(x,y,t)]\, b(x,y,t) \qquad (1)$$

where f(x,y,t) is the recorded sequence, o(x,y,t) is the moving objects function in the sequence, $O_m(x,y,t)$ is a binary valued mask function which determines the target presence (it is 1 for all pixels corresponding to the moving object and is zero otherwise), and b(x,y,t) is the time-varying background.

In this model, the moving objects occlude the background which is the case in real world scenes. In this work we will address the moving objects term, i.e o(x,y,t). The effects of image background on the velocity computation of moving objects using the above model is addressed in [35].

To simplify the analysis of the sequence, the projections of the sequence on the x- and y-axis are processed [36]. The use of projections in motion analysis is efficient. The projections of images destroy the structure of the objects in the image sequence. However, it preserves the information necessary for motion estimation. The projections of the sequence in the x- and y- directions are given by:

$$f(x,t) = \sum_{y=0}^{y=M-1} f(x,y,t), \qquad (2.a)$$

$$f(y,t) = \sum_{x=0}^{x=N-1} f(x,y,t), \qquad (2.b)$$

where N and M are the number of pixels of the image in the x and y-directions, respectively. The x- and y-directions are along the axis of the images. The first horizontal

scan line of every image is taken as the x-axis and the vertical line of pixels corresponding to the first location of each scan line as the y-axis.

The two-dimensional sequence is transformed to two one-dimensional sequences. The sequences f(x,t) and f(y,t) are processed and the velocities V_x and V_y of the objects are computed. The velocity (V_i) of moving object i, in the f(x,y,t) sequence is computed from :

$$[V_i] = [V_{ix} \quad V_{iy}]^T \tag{3}$$

A model for multiple moving objects in one-dimensional time sequence, each of one pixel in size, is given by :

$$o[n,m] = \sum_{i=1}^{i=r} A_i \delta[n-L_i-mV_i] \prod_{j=1}^{j=i-1} (1-\delta[n-L_j-mV_j]) \tag{4}$$

Where n,m are the pixel co-ordinates at location n and frame m, r is the number of moving objects in the sequence, A_i is the grey level for moving object i (i=1,r), L_i is the initial distance position (from the first location of the first frame) of moving object i, V_i is the velocity of moving object i, and $\delta[\ ]$ is the dirac-delta function. The use of moving objects which are one pixel in size possess no limitation as large moving objects may be looked at as a number of moving objects with identical velocity that are adjacent in location. The applicability of this model to the case of large moving objects is addressed in[36].

The Discrete Sine Area Transform(DSAT) of equation (4) is given by:

$$S[k,f] = \sum_{n=0}^{n=N-1} \sum_{m=0}^{m=M-1} o[n,m] \sin[2\pi kn/N] \sin[2\pi fm/M] \tag{5}$$

where N is the number of pixels in the image, and M is the number of frames in the image sequence. Applying the DSAT to the model of Equation (4), the result of the formulation is given by the following equation (For detailed formulations reference may be made to the appendix) :

$$S[k,f] = \frac{1}{2}\sum_{i=1}^{i=r} M_i A_i \{ \frac{\text{sinc}[kV_i-f)M_i/N]}{\text{sinc}[(kV_i-f)/N]} \cos[\pi\{2kL_i+(kV_i-f)(2m_i+M_i-1)\}/N]$$

$$- \frac{\text{sinc}[kV_i+f)M_i/N]}{\text{sinc}[(kV_i+f)/N]} \cos[\pi\{2kL_i+(kV_i+f)(2m_i+M_i-1)\}/N]$$

$$- \frac{1}{2}\sum_{i=1}^{i=r}\sum_{j=1}^{j=i-1} A_i \{\cos[2\pi\{kL_i+(kV_i-f)(L_j-L_i)/(V_i-V_j)\}/N] - \cos[2\pi\{kL_i+(kV_i+f)(L_j-L_i)/(V_i-V_j)\}/N]$$

$$\varepsilon[m_j,(L_j-L_i)/(V_i-V_j),M_{ij}]$$

$$+ \frac{1}{2}\sum_{i=1}^{i=r}\sum_{j=1}^{j=i-1}\sum_{p=1}^{p=j-1} A_i \delta[L_p-L_i+(V_i-V_p)(L_j-L_i)/(V_i-V_j)]$$

$$\{\cos[2\pi\{kL_i+(kV_i-f)(L_j-L_i)/(V_i-V_j)\}/N]$$

$$-\cos[2\pi\{kL_i+(kV_i+f)(L_j-L_i)/(V_i-V_j)\}/N]\}\varepsilon[m_p,(L_j-L_i)/(V_i-V_j,M_{ijp}] + \ldots \quad (6)$$

where ε is as defined in Equation (A.6) of the appendix.

It is clear from Equation (6) that the contribution of the moving objects is proportional to the amplitude of the moving objects (A_i) and to the number of frames the moving objects are in the sequence (M_i). The effect of the occlusion terms is very small such that it has no appreciable effect on the result. Hence, equation (6) can be approximated by the first two terms. Each of the object terms of equation(6) is nonzero only at $f=\pm kV_i$. This indicates that the locations of the spectral peaks are related to the velocities of the moving objects. Hence the velocities of the moving objects are found by detecting peaks in the spectrum given by equation(6) and finding the temporal frequencies at which the peaks exist. The velocities of the moving objects are given by :

$$V_i = \pm f_{ip} / k_s \quad (7)$$

where $f_{1p}, f_{2p}, \ldots, f_{rp}$ are the temporal frequencies of the peaks, and k_s is a specified spatial frequency.

A simple peak detection algorithm is used to detect the peaks of the Sine spectrum. This algorithm scans the Fourier spectrum at the selected spatial frequency corresponding to the required resolution. The spectrum at zero temporal frequency is skipped as it corresponds to stationary objects. Other spectral values are compared to detect the peak value. The temporal frequencies corresponding to the detected peaks are found. The velocities of the moving objects are found by dividing the temporal frequencies f_i's by the respective spatial frequencies k's. It is clear from equations (6 and 7) that the spectrum resulting from a moving object has two components, one component in the positive frequency range and the other in the negative frequency range. Hence, the direction of motion is not unique. This technique computes the amplitudes of the velocities of the moving objects but the direction is ambiguous. To compute the directions, the Fourier spectral values corresponding to the temporal and spatial frequencies of the SAT spectral peaks are computed (i.e, $F[f_{ip}, k_s]$ and $F[-f_{ip}, k_s]$). If $F[f_{ip}, k_s]$ is greater than $F[-f_{ip}, k_s]$ then the direction of motion is opposite to that of f_{ip} ; otherwise, it has the same direction as f_{ip} as was shown in [36]. Alternatively, the technique of [37] may be used to estimate the direction of the moving objects. An interesting result of Equation (6) is that the contribution of stationary objects is zero. Hence, stationary objects do not affect this technique. This is an additional advantage, as stationary objects (in other techniques) make the image analysis more complex and requires extra processing time.

This algorithm is more efficient compared with the Fourier transform used in spatio-temporal frequency techniques. It has been shown by Hou[31] that the number of arithmetic operations in computing the FFT for complex data is nearly double that for real data. Computing DSAT requires the use of a real FFT (as only the imaginary part of the Fourier transform is required to compute DSAT) while the Fourier analysis of two-dimensional sequence requires the use of one real and one complex FFTs, hence, using DSAT is faster. In addition, in the case of DSAT it is enough to process half of the spectrum (either the positive or the negative frequency spectrum) to compute the velocities. Therefore, an additional reduction in processing time is achieved. The output of DSAT is real while the output in the case of Fourier analysis is complex. Motion analysis using Fourier spectrum [14,36] requires the computation of the amplitude of the spectrum which requires an additional $2N^2$ multiplications and N^2 additions (in the two-dimensional case), not counting the square root computation. Since two-dimensional processing is used, the reduction in processing time is considerable. In addition, one two-dimensional image array is needed using this technique while two two dimensional image arrays(real and imaginary) are needed in the case of Fourier analysis. This results in considerable saving in memory requirements, especially when using personal computers with limited memory.

Experimental Results

The simulations were carried out on an IBM PS/2 model 80 computer using C programming language. The graphs were produced using Surfer Access System software of Golden Software Inc.

Figure 1(a) shows an image sequence with 4 moving objects projected on the x-axis(i.e the x-sequence). The velocities of the four moving objects are 0.5,1,-1,and -2 pixels per frame(ppf). The moving objects initially started at different locations 24,-16,59, and 56, respectively. The figure shows the paths of the moving objects and the occlusions between the objects when any two or more objects are at the same location and at the same frame. Figure 1(b) shows the SAT spectrum of the time sequence of Fig. 1(a). The figure shows the spectral peaks corresponding to the moving objects. Each moving object has two spectral peaks at two different temporal frequencies(one at positive and another at negative frequencies) for each spatial frequency. The temporal frequencies are related to the amplitude of the velocities of the moving objects as given by equation (7). The spectral peaks for each object are shifting according to the value of the spatial frequency, in accordance with equation (7). Figure 1(c) shows the DSAT spectrum of Fig. 1(a) at a spatial frequency of 4.

It is clear from the figure that there are 6 peaks, three in the positive frequency range and three in the negative frequency range. The peaks correspond to temporal frequencies of $\pm 2, \pm 4$, and ± 8. This indicates that the amplitudes of the velocities of the moving objects are 2/4=0.5, 4/4=1, and 8/4=2 ppf, respectively. To estimate the direction of motion the Fourier spectrum at a spatial frequency of four and temporal frequencies of $\pm 2, \pm 4$, and ± 8 are computed. The spectrum values at both sides of the zero temporal frequency (for example, at temporal frequencies of 2 and -2) are compared. The direction of motion is opposite to the direction of the temporal frequency of the peak in the Fourier spectrum [36]. The Fourier spectrum for the above locations are given in the table below.

Comparing the spectral amplitudes it is clear that the spectrum at frequencies -2,-4,4, and 8 are the peaks. This indicates that the sign of the velocities are positive for the first and second objects, and negative for the others. Hence, the velocities of the moving objects are 0.5,1,-1, and -2 ppf. We have two peaks at temporal frequencies of - 4 and 4 (which are close in amplitude) because the contributions of objects two and three(with velocities of -1 and 1) have their spectral contributions at temporal frequencies of - 4 and 4. This case indicates the presence of two moving objects with same velocity amplitude and opposite directions. Figure 1(d) shows the Fourier spectrum at a spatial frequency of 4. The figure has four peaks, two corresponding to positive temporal frequencies and the other peaks are corresponding to negative temporal frequencies. Hence, indicating the

directions of the velocities of the moving objects as shown above. The little peak corresponding to zero temporal frequency is due to the dc component of the Fourier spectrum.

Temporal frequency f	-2	2	- 4	4	- 8	8
|F[4, f]|	427.7	53.4	326	349	7.9	473.3

Figure 2(a) shows the same sequence of Fig. 1(a) but with an additional stationary object of 5 pixels in width. Fig. 2(b) shows the DSAT spectrum of the sequence in Fig. 2(a). Figure 2(c) shows the DSAT spectrum at a spatial frequency of 4. Figures 2(b) and 2(c) are similar to Figs. 1(b) and 1(c), respectively. It is clear from these figures that the presence of stationary objects does not affect the DSAT spectrum of the sequence and hence does not affect the presented technique. Figs. 3(a) and 3(b) show the Fourier spectrum of the sequences of Figs. 1(a) and 2(a), respectively. Comparing these two figures and Figs.1(d) and 2(d), it is clear that the presence of stationary objects increase the Fourier spectrum values at spatial and temporal frequencies of 0. Hence, the Sine Area Transform, which is not affected by stationary objects, is better in this respect. The amplitudes of the peaks in the Fourier spectrum is approximately double the peak values in the SAT spectrum, thus confirming the analytical results of section II.

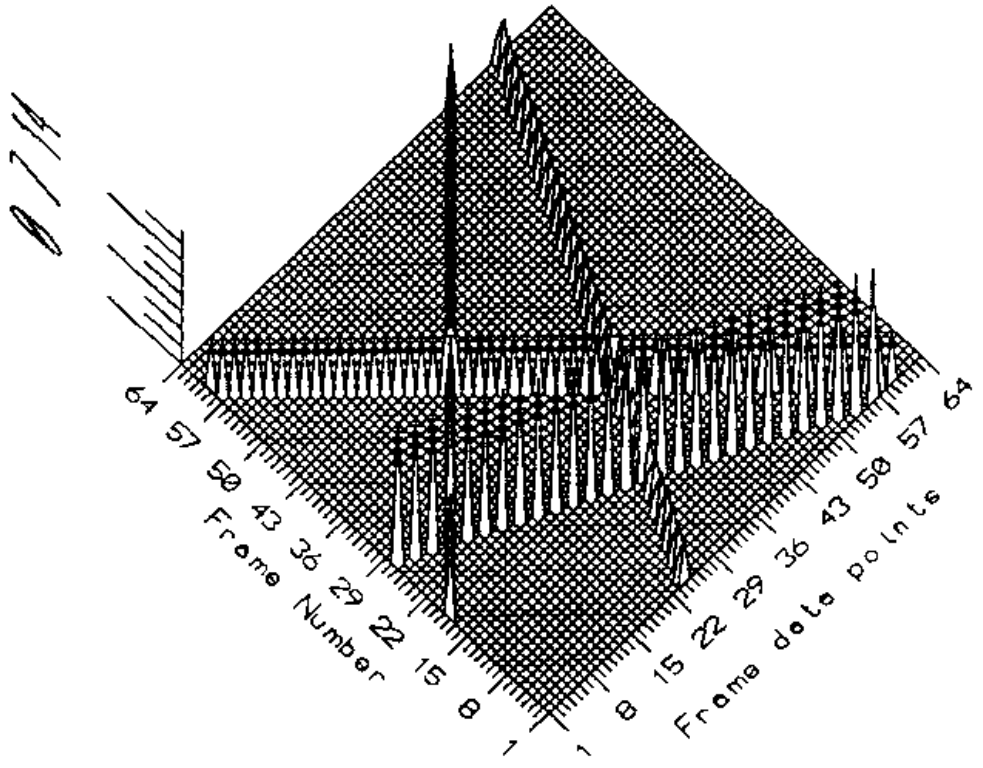

Fig. 1a. X-sequence with 4 moving objects

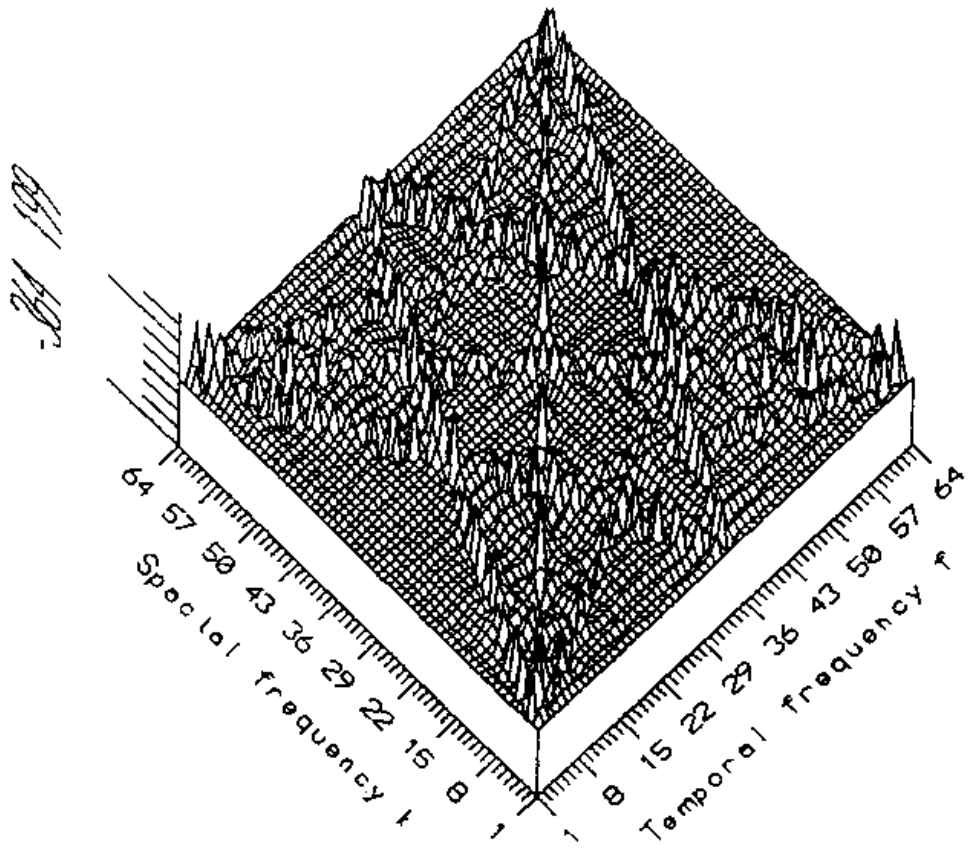

Fig. 1b. SAT spectrum of the sequence of Fig. 1a

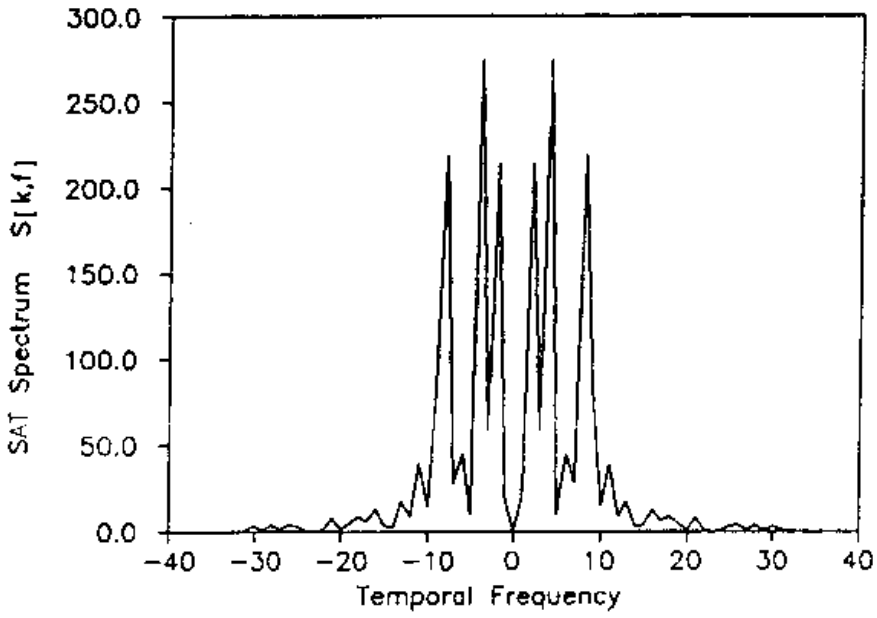

Fig. 1c. SAT spectrum of Fig. 1b at special frequency of 4

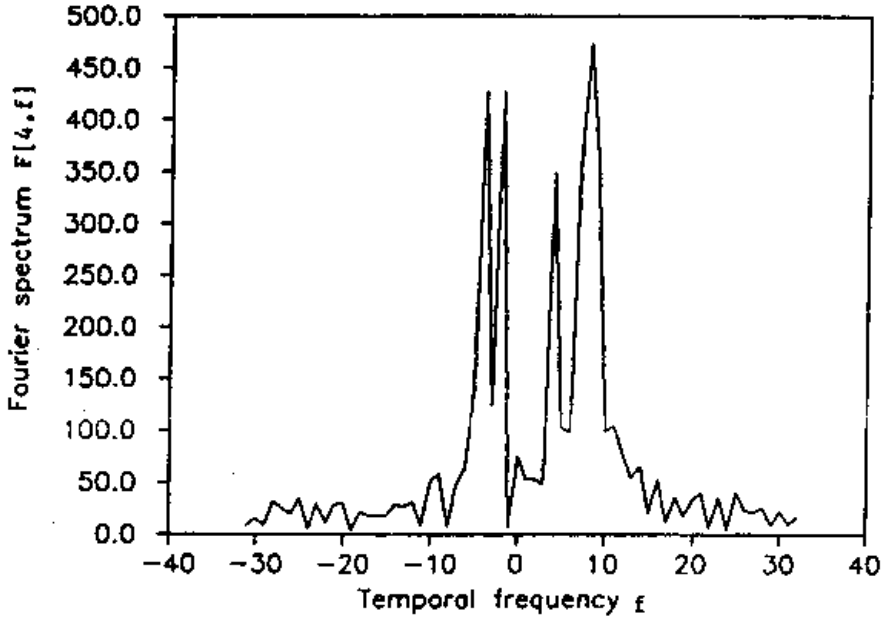

Fig. 1d. Fourier spectrum of Fig. 1a at a special frequency of 4

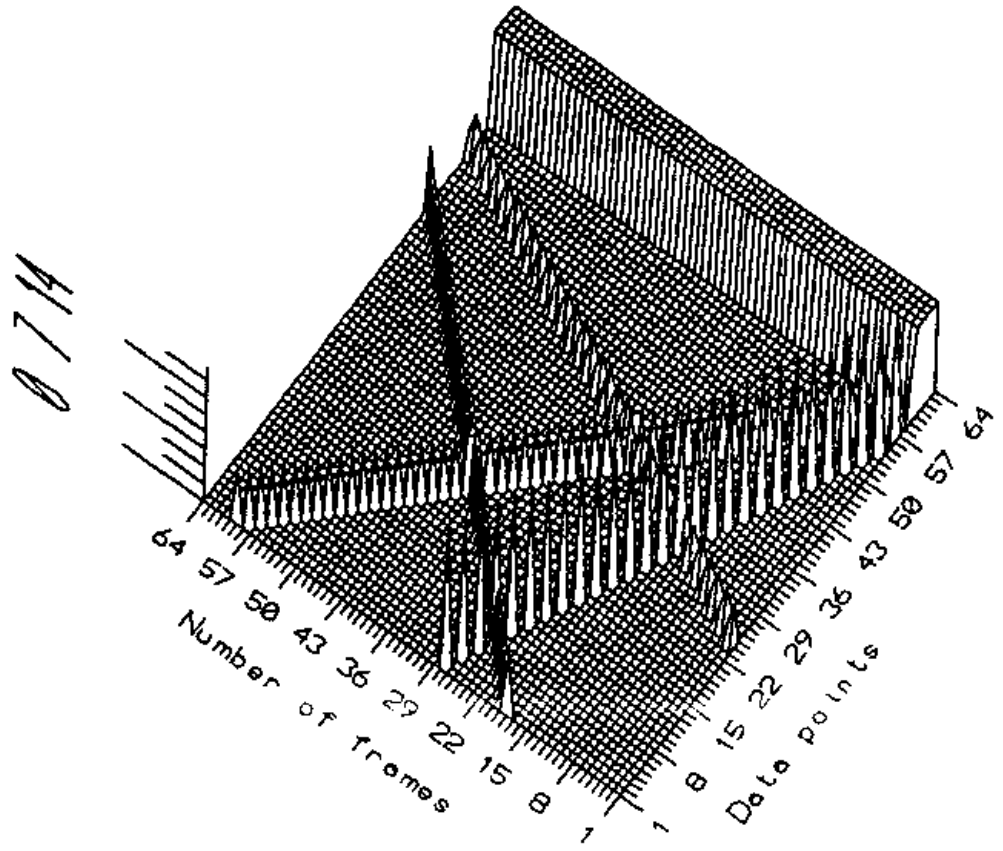

Fig. 2a. X-sequence of Fig. 1a with stationary object .

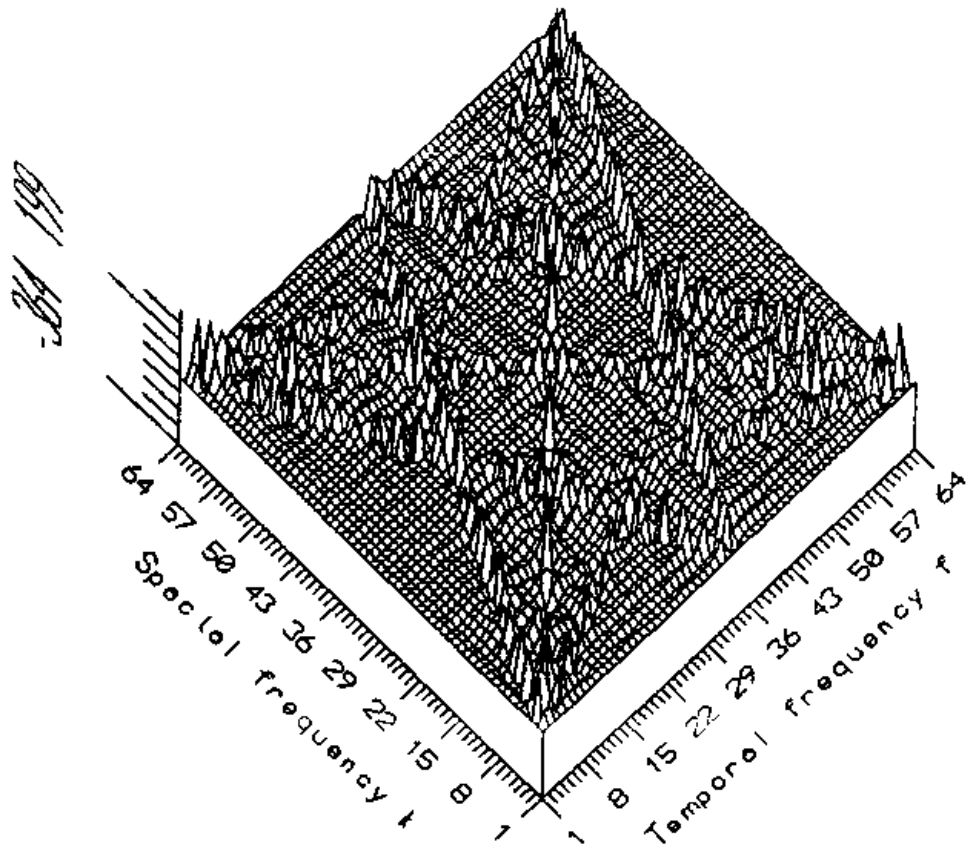

Fig. 2b. SAT spectrum of the sequence of Fig. 2a

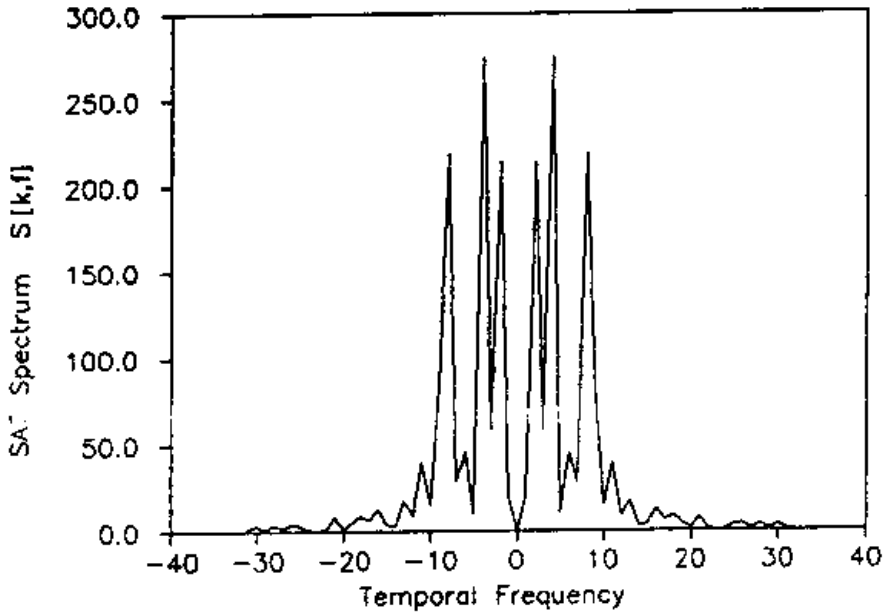

Fig. 2c. SAT spectrum of Fig. 2b at a special frequency of 4

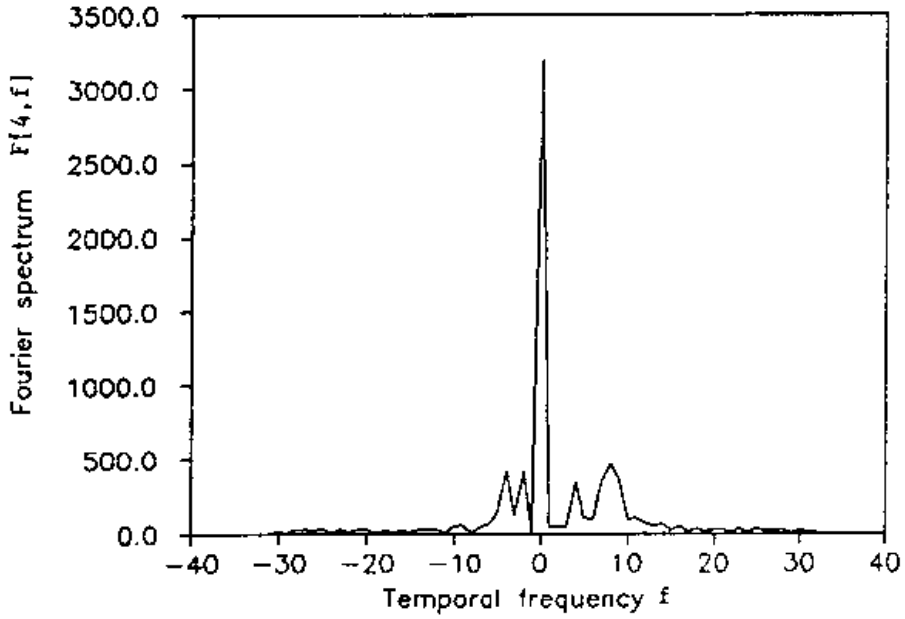

Fig. 2d. Fourier spectrum of sequence of Fig. 2a at a speial frequency of 4

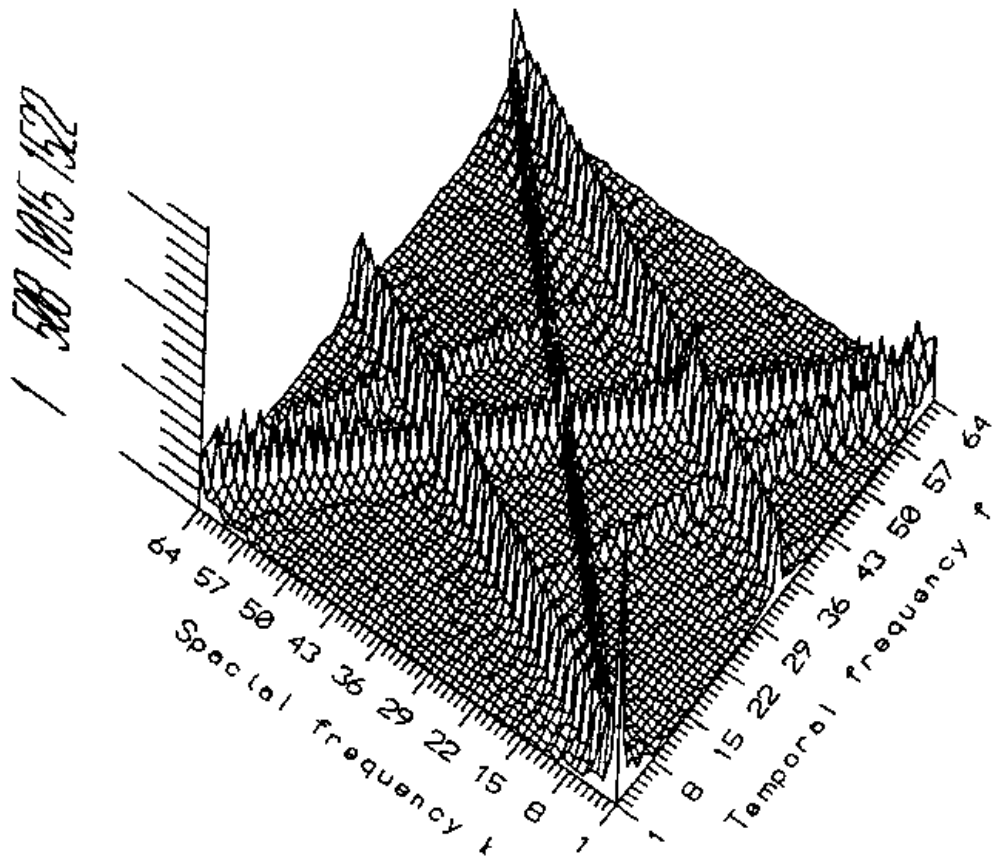

Fig. 3a. Fourier spectrum of Fig. 1a

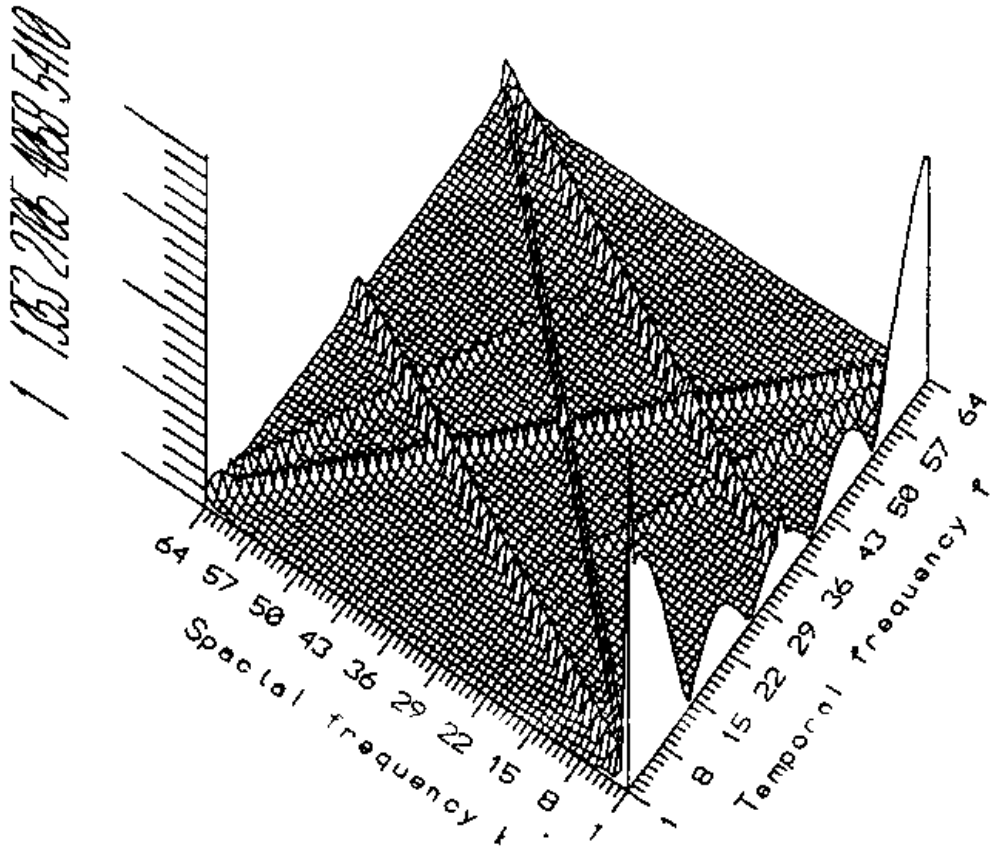

Fig. 3b. Fourier spectrum of Fig. 2a

Conclusion

The proposed algorithm gives an accurate estimate of the velocities of moving objects. Using an image sequence and not two or three frames only (as is the case with other techniques) gives a better estimate of the velocity in the sequence. The algorithm is fast as a real FFT is used in the computation of the DSAT and only a simple peak detection algorithm is needed. The analysis is applied to half the spectrum only (the spectrum corresponding to the positive or to the negative frequencies only), hence a further reduction in computations.

The presence of stationary objects does not change the DSAT spectral values. This is not the case with the Fourier spectrum at spatial and temporal frequencies of 0 which increases due to the presence of stationary objects. Hence, the presence of stationary objects does not cause the technique to be more complex or to require higher processing time to distinguish stationary- from moving-objects. The use of this technique does not require the segmentation or matching of the static images nor finding the correspondence

between frames. This technique, however, requires a number of frames which are hard to obtain in some applications.

The applicability of this technique has been demonstrated by a new mathematical formulation and experimental results obtained on PS/2.

Acknowledgement. The author would like to express his gratitude to the referees for their time and effort in carefully reviewing the preliminary manuscript and for their constructive criticism and for bringing reference [23] to his attention. The incorporation of their suggestions and comments significantly improved the clarity of the final manuscript.

Appendix

Derivation of the dsat for a sequence with multiple-moving objects.

A model for multiple moving objects in one-dimensional time sequence, each of one pixel in size, is given by :

$$o[n,m] = \sum_{i=1}^{i=r} A_i \delta[n - L_i - mV_i] - \sum_{i=1}^{i=r} \sum_{j=1}^{j=i-1} A_i \delta[n - L_i - mV_i]\delta[n - L_j - mV_j]$$

$$+ \sum_{i=1}^{i=r} \sum_{j=1}^{j=i-1} \sum_{p=1}^{p=j-1} A_i \delta[n - L_i - mV_i]\delta[n - L_j - mV_j]\delta[n - L_p - mV_p]$$

$$- \sum_{i=1}^{i=r} \sum_{j=1}^{j=i-1} \sum_{p=1}^{p=j-1} \sum_{q=1}^{q=p-1} A_i \delta[n - L_i - mV_i]\delta[n - L_j - mV_j]\delta[n - L_p - mV_p]\delta[n - L_q - mV_q] + .. \quad \text{(A.1)}$$

where n,m are the pixel co-ordinates at data point n & frame m, r is the number of moving objects in the sequence, A_i is the grey level for moving object i (i=1,r), L_i is the initial distance position (from the first location of the first frame) of moving object i, V_i is the velocity of moving object i, and δ [] is the dirac-delta function.

The above model is an expansion of the model given in equation(4). In this model, we arbitrarily assumed that the lower indexed moving object occludes the higher indexed moving object when they cross in the sequence. This occurs when the moving objects are at the same location and at the same frame (when other rules are applied, only the amplitudes of the occlusion terms are changed accordingly).

The first summation term represents the trajectories of the individual moving objects; the second summation represents the occlusion terms, when one moving object occludes another in the sequence; the third summation represents the occlusion, when one moving object occludes two moving objects; other summation terms are added to represent a moving object occluding higher number of objects

The Discrete Sine Area Transform(DSAT) of equation (A.1) is given by:

$$S[K,f] = \sum_{n=0}^{n=N-1} \sum_{m=0}^{m=M-1} o[n,m] \sin[2\pi kn/N] \sin[2\pi fm/M] \quad \text{(A.2)}$$

where N is the number of pixels in the image, M is the number of frames in the image sequence, and $\tau = 2\pi/N$.

The moving objects appear in the sequence in frames $[m_i : m_i + M_i - 1]$ for moving object i,(i=1,r); frames $[m_i : M_{ij} - 1]$ for the occlusion term of two objects, (i=1,r and j=1,i-1); frames $[m_p : M_{ijp} - 1]$ for the occlusion term of three objects, (i=1,r,j=1,i-1, and p=1,j-1); and other frames are added similarly.

where M_i is the initial frame number at which moving object i entered the sequence; M_i is the number of frames in which moving object i is in the sequence; $M_{ij} = \min[(m_i + M_i),(m_j + M_j)]$, is the last frame in which moving objects i and j are simultaneously in the sequence;

$M_{ijp} = \min[(m_i + M_i),(m_j + M_j),(m_p + M_p)]$, is the last frame in which i,j,p moving objects are simultanously in the sequence.; and other terms are defined similarly.

Applying the limits of the trajectories of the moving objects in the image sequence (since the moving objects are not present in the images outside these limits), we have

$$S[K,f] = \sum_{i=1}^{i=r} \sum_{m=m_i}^{m=m_i+M_{i-1}} \sum_{n=0}^{n=N-1} A_i \delta[n-L_i-mV_i] \sin[\tau kn] \sin[\tau fm]$$

$$- \sum_{i=1}^{i=r} \sum_{j=1}^{j=i-1} \sum_{m=m_j}^{m=M_{ij}-1} \sum_{n=0}^{n=N-1} A_i \delta[n-L_i-mV_i]\delta[n-L_j-mV_j] \sin(\tau kn)\sin(\tau fm)$$

$$+ \sum_{i=1}^{i=r} \sum_{j=1}^{j=i-1} \sum_{p=1}^{p=j-1} \sum_{m=m_i}^{m=M_{ijp}-1} \sum_{n=0}^{n=N-1} A_i \delta[n-L_i-mV_i]\delta[n-L_j-mV_j]\, \delta[n-L_p-mV_p]$$

$$\sin(\tau kn)\, \sin(\tau fm)$$

$$- \sum_{i=1}^{i=r} \sum_{j=1}^{j=i-1} \sum_{p=1}^{p=j-1} \sum_{q=1}^{q=p-1} \sum_{m=m_i}^{m=M_{ijpq}-1} \sum_{n=0}^{n=N-1} A_i \delta[n-L_i-mV_i]\delta[n-L_j-mV_j]\, x$$

$$\delta[n-L_p-mV_p]\delta[n-L_q-mV_q] \sin(\tau kn)\sin(\tau fm) + \ldots \qquad (A.3)$$

Summing with respect to n, all the terms are zero except those at n between 0 and N-1

and $n=L_i+mV_i$ for the first term,(where i=1,r), $n=L_i+mV_i$ and $n=L_j+mV_j$ for the second term,(where i=1,r and j=1,i-1), and so on. The result of the summation with respect to n is given by:

$$S[K,f] = \sum_{i=1}^{i=r} \sum_{m=m_i}^{m=m_i+M_i-1} \sin[\tau k(L_i+mV_i] \sin[\tau fm]$$

$$-\sum_{i=1}^{i=r} \sum_{j=1}^{j=i-1} \sum_{m=m_j}^{m=M_{ij}-1} A_i\delta[L_i-L_j+m(V_i-V_j)] \sin(\tau k(L_i+mV_i))\sin(\tau fm)$$

$$+\sum_{i=1}^{i=r} \sum_{j=1}^{j=i-1} \sum_{p=1}^{p=j-1} \sum_{m=m_i}^{m=M_{ijp}-1} A_i\delta[L_i-L_j+m(V_i-V_j)]\delta[L_i-L_p+m(V_i-V_p)]\sin(\tau k(L_i+mV_i)) \sin(\tau fm)$$

$$-\sum_{i=1}^{i=r} \sum_{j=1}^{j=i-1} \sum_{p=1}^{p=j-1} \sum_{q=1}^{q=p-1} \sum_{m=m_i}^{m=M_{ijpq}-1} A_i\delta[L_i-L_j+m(V_i-V_j)] \text{ x}$$

$$\delta[L_i-L_p+m(V_i-V_p)]\delta[L_i-L_q+m(V_i-V_q)] \sin(\tau k(n)\sin(\tau fL_i+mV_i)) \sin(\tau fm) +\ldots \quad \text{(A.4).}$$

The first term is given by:

$$\sum_{i=1}^{i=r} \sum_{m=m_i}^{m=m_i+M_i-1} A_i \sin[\tau k(L_i+mV_i] \sin[\tau fm]$$

Summing the first term with respect to m, we have :

$$=\frac{1}{2}\sum_{i=1}^{i=r} \sum_{m=m_i}^{m=m_i+M_i-1} A_i \{\cos[\tau\{m(kV_i-f)+kL_i\}]- \cos[\tau\{m(kV_i+f)+kL_i\}]\}$$

$$= \frac{1}{4}\sum_{i=1}^{i=r}\sum_{m=m_i}^{m=m_i+M_i-1} A_i \left\{ \frac{e^{j\tau[m(kV_i-f)+kL_i]} + e^{-j\tau[m(kV_i-f)+kL_i]}}{e^{j\tau[m(kV_i-f)+kL_i]} - e^{-j\tau[m(kV_i-f)+kL_i]}} \right\}$$

$$= \frac{1}{4}\sum_{i=1}^{i=r} M_i A_i \left\{ \begin{array}{l} e^{\frac{j\tau[kL_i+(kV_i+f)(2m_i+M_i-1)/2]}{\text{sinc}[\tau(kV_i-f)/2\pi]}\text{sinc}[\tau(kV_i-f)M_i/2\pi]} + e^{\frac{-j\tau[kL_i+(kV_i-f)(2m_i+M_i-1)/2]}{\text{sinc}[\tau(kV_i-f)/2\pi]}\text{sinc}[\tau(kV_i-f)M_i/2\pi]} \\ -e^{\frac{j\tau[kL_i+(kV_i+f)(2m_i+M_i-1)/2]}{\text{sinc}[\tau(kV_i+f)/2\pi]}\text{sinc}[\tau(kV_i+f)M_i/2\pi]} - e^{\frac{-j\tau[kL_i+(kV_i+f)(2m_i+M_i-1)/2]}{\text{sinc}[\tau(kV_i+f)/2\pi]}\text{sinc}[\tau(kV_i+f)M_i/2\pi]} \end{array} \right\}$$

Combining the exponential terms and substituting the value of τ, we have

$$= \frac{1}{2}\sum_{i=1}^{i=r} M_i A_i \left\{ \begin{array}{l} \frac{\text{sinc}[(kV_i-f)M_i/N]}{\text{sinc}[(kV_i-f)/N]}\cos[\pi\{2KL_i+(kV_i-F)(2m_i+M_i-1)\}/N] \\ \frac{-\text{sinc}[(kV_i+f)M_i/N]}{\text{sinc}[(kV_i+f)/N]}\cos[\pi\{2KL_i+(kV_i+F)(2m_i+M_i-1)\}/N] \end{array} \right\} \quad (A.5)$$

where

$$\sum_{m=m_i}^{m=m_i+M_i-1} e^{j\tau(m(kV_i+f)} = \frac{e^{j\tau(kV_i+f)m_i} - e^{j\tau(kV_i+f)(m_i+M_i}}{1 - e^{j\tau(kV_i+f)}}$$

$$= e^{j\tau(kV_i+f)[m_i+M_i-1)/2]}\frac{e^{j\tau(kV_i+f)M_i/2} - e^{-j\tau(kV_i+f)M_i/2}}{e^{j\tau(kV_i+f)/2} - e^{-j\tau(kV_i+f)/2}} \quad (A.6)$$

$$= M_i e^{j\tau(kV_i+f)[m_i+M_i-1)/2]}\frac{\text{sinc}[\tau(kV_i+f)M_i/2\pi}{\text{sinc}[\tau(kV_i+f)/2\pi}$$

The second term :

Let ε be a function such that $\varepsilon[x1,x,x2] = 1$ for $x1 \leq x < x2$

$= 0$ otherwise (A.7)

or , $\varepsilon[x1,x,x2] = u[x-x1] - u[x-x2]$, where $u[x]$ is the unit step function.

This function is used to constrain the trajectories of the moving objects within image limits. Summing with respect to m, the second term is zero except at m between the limits of the summation and

$$m=(L_j - L_i)/V_i - V_j \text{ (the condition is } L_i - L_j + m(V_i - V_j)=0).$$

The conditions for the second term (to be nonzero) is defined in terms of the ε function, as $\varepsilon[m_j,(L_j - L_i)/(V_i - V_j),M_{ij}]$, where i=1, r and j=1,i-1.

$$-\sum_{i=1}^{i=r}\sum_{j=1}^{j=i-1}\sum_{m=m_j}^{m=M_{ij}-1} A_i\delta[L_j-L_i+m(V_i - V_j] \sin(\tau k(L_i+mV_i)) \sin(\tau fm)$$

$$=\sum_{i=1}^{i=r}\sum_{j=1}^{j=i-1} A_i \sin[\tau k\{L_i+V_i(L_j-L_i)/(V_i-V_j)\}] \sin[\tau f(L_j-L_i)/(L_i-L_j)]\varepsilon[m_j,(L_j-L_i)/(V_i-V_j),M_{ij}].$$

$$= -\sum_{i=1}^{i=r}\sum_{j=1}^{j=i-1} A_i \{\cos[\tau\{kL_i+(kV_i-f)(L_j - L_i)/(V_i-V_j)\}]$$

$$- \cos[\tau\{kL_i+(kV_i+f)(L_j - L_i)/(V_i-V_j)\}] \varepsilon[m_j,(L_j - L_i)/(V_i-V_j),M_{ij}].$$

(A.8)

The third term :

Summing with respect to m, the third term is zero except at m between the limits of the summation and $m=(L_j-L_i)/(V_i-V_j)$ and $m=(L_p-L_i)/(V_i-V_p)$ (or $L_i-L_j+m(V_i-V_j=0$ and $L_i-L_p+m(V_i-V_p)=0$). It is defined in terms of the ε function as $\varepsilon[m_p,(L_j-L_i)/(V_i-V_j),M_{ijp}]$.

$$\sum_{i=1}^{i=r}\sum_{j=1}^{j=i-1}\sum_{p=1}^{p=j-1}\sum_{m=m_i}^{m=M_{ijp}-1} A_i d[L_i-L_j+m(V_i-V_j)]d[L_i-L_p+m(V_i-V_p)]\sin(tk(L_i+mV_i)) \sin(tfm)$$

$$=\sum_{i=1}^{i=r}\sum_{j=1}^{j=i-1}\sum_{p=1}^{p=j-1}A_i\delta[L_i-L_p+(V_i-V_p)(L_j-L_i)/(V_i-V_j)]\sin(\tau k(L_i+V_i(L_j-L_i)/(V_i-V_j)))\ \sin(\tau f(L_j-L_i)/(V_i-V_j))$$

$$=\frac{1}{2}\sum_{i=1}^{i=r}\sum_{j=1}^{j=i-1}\sum_{p=1}^{p=j-1}A_i\delta[L_i-L_p+(V_i-V_p)(L_j-L_i)/(V_i-V_j)]\cos(\tau(kL_i+(kV_i-f)(L_j-L_i)/(V_i-V_j)))$$
$$-\cos(\tau(kL_i+(kV_i-f)(L_j-L_i)/(V_i-V_j)))\}\ \varepsilon[m_p,(L_j-L_i)/(V_i-V_j),M_{ijp}].$$

(A.9)

Other terms may be added similarly.

Substituting the results of summing the above terms and substituting the value of τ, we have

$$S[k,f]=\frac{1}{2}\sum_{i=1}^{i=r}M_i\,A_i\left\{\begin{array}{l}\dfrac{\mathrm{sinc}[(kV_i-f)\,M_i/N]}{\mathrm{sinc}[(kV_i-f)/N]}\cos[\pi\{2kL_i+(kV_i-F)(2m_i+M_i-1)\}/N]\\ \dfrac{-\mathrm{sinc}[(kV_i+f)\,M_i/N]}{\mathrm{sinc}[(kV_i+f)/N]}\cos[\pi\{2kL_i+(kV_i+F)(2m_i+M_i-1)\}/N]\end{array}\right\}$$

$$-\sum_{i=1}^{i=r}\sum_{j=1}^{j=i-1}A_i\{\cos[2\pi\{kL_i+(kV_i-f)(L_j-L_i)/(V_i-V_j)\}/N]$$
$$-\cos[2\pi\{kL_i+(kV_i+f)(L_j-L_i)/(V_i-V_j)\}/N\ \varepsilon[m_j,(L_j-L_i)/(V_i-V_j),M_{ij}]$$

$$+\sum_{i=1}^{i=r}\sum_{j=1}^{j=i-1}\sum_{p=1}^{p=j-1}A_i\,\delta[L_p-L_i+(V_i-V_p)(L_j-L_i)/(V_i-V_j]\{\cos[2\pi\{kL_i+(kV_i-f)(L_j-L_i)/(V_i-V_j)\}/N]$$
$$-\cos[2\pi\{kL_i+(kV_i+f)(L_j-L_i)/(V_i-V_j)\}/N\ \varepsilon[m_p,(L_j-L_i)/(V_i-V_j),M_{ijp}]+\ldots$$

(A.10)

References

[1] Cafforio, C. and Fabio, R. "Methods for Measuring Small Displacements of TV Images." *IEEE Transactions on Information Theory*, IT-22, No.5 (Sept 1976), 573-79.

[2] Fennema, C.L. and Thompson, W.B. "Velocity Determination in Scenes Contain-ing Several Moving Objects." *Computer Graphics and Image Processing*, 9, (April 1979), 301-15 .

[3] Nagel, H.H and Enkelmann,W. "An Investigation of the Smoothness Constraints for the Estimation of Displacement Vector Fields from Image Sequences." *IEEE Transactions on PAMI*, PAMI-8, No.5 (1986), 565-93.

[4] Kearney, J.K. *et al.* "Optical Flow Estimation: An Error Analysis of Gradient-based Methods with Local Optimization." *IEEE Transactions on PAMI*, PAMI 9, No.2 (1987), 229-44.

[5] Jain, R. and Nagel, H. H. "On the Analysis of Accumulative Difference Pictures from Image Sequences of Real World Scenes." *IEEE Transactions on PAMI*, PAMI-1, No2 (April 1979), 206-14.

[6] Jain, R. ; Martin,W.N. and Aggarwal, J.K. "Segmentation Through the Detection of Changes Due to Motion." *Computer Graphics and Image Processing*,11,(Sept. 1979), 13-34.

[7] Jain, R. "Extraction of Motion Information from Peripheral Processing." *IEEE Transactions on PAMI*, 3, No.5, (1981), 489-503.

[8] Aggrawal, J.K. and Duda, R.O. " Computer Analysis of Moving Polygonal Images." *IEEE Transactions on Computers*, C-24, No.10 (Oct 1975), 966-975.

[9] Martin,W.N. and Aggrawal, J.K."Computer Analysis of Dynamic Scenes Contain-ing Curvilinear Figures." *Pattern Recognition* ,11 (1979), 169-78.

[10] Murray, D.W. and Buxton, B.F. "Scene Segmentation from Visual Motion Using Global Optimization."*IEEE Transactions on PAMI*, 9 , No.2 (1987), 220-28.

[11] Peleg, S. and Rom, H."Motion Based Segmentation." *Proceedings of the Interna-tional Conference on Pattern Recognition,* Atlantic City, USA,1 (June 1990),103 -13.

[12] Girod, B. and Kuo, D. "Direct Estimation of Displacement Histograms." *Proc. Image Understanding Machine Vision*, Optical Soc. of America (Cape Cod), June 1989, pp.73-76.

[13] Barnard, S.T.and Thompson,W.B."Disparity Analysis of Images."*IEEE Transac-tions on PAMI*, PAMI-2, No.4, (1980), 333-340.

[14] Jacobson, L. and Wechsler, H. "Derivation of Optical Flow Using a Spatio-temporal-frequency Approach." *Computer Vision, Graphics, and Image Processing*, 38 (1987), 29-65.

[15] Horn, B.K.P. and Schunk, B.G. "Determining Optical Flow." *Artificial Intellig-ence*, 17 (May 1981), 185-203.

[16] Schunk, B.G. "The Image Flow Constraint." *Computer Vision, Graphics, and Image Processing*, 35 (1986), 20-46.

[17] Adiv, G. "Inherent Ambiguities in Recovering 3-D motion and Structure from Noisy Flow Field." *IEEE Transactions on PAMI*, 11, No.5 (May 1989), 477-89.

[18] Verri, A. and Poggio,T. "Motion Field and Optical Flow : Qualitative Properties." *IEEE Transactions on PAMI*, 11, No.5 (May 1989), 490-98.

[19] Aisbett, J. "Optical Flow with an Intensity-weighted Smoothing." *IEEE Transac-tions on PAMI*, 11, No.5 (May 1989), 512-22.

[20] Heeger, David J. "Optical Flow from Spatio-temporal Filters". *Proceedings of the First International Conference on Computer Vision*, London, England, (June 1987), 181-90 .

[21] Bergen, J.R. *et al.* "A Three-frame Algorithm for Estimating Two-component Image Motion."*IEEE Transactions on PAMI* , 14,No.9 (1992), 869-85.

[22] Wohn, Kwang Y. *et al.* " A Contour-based Recovery of Image Flow: Interactive Transformation Method". *IEEE Transactions on PAMI*, 13, No.8 (1992), 746-60.

[23] Young, R.W. and Kingsbury, N.G. " Frequency Domain Motion Estimation Using a Complex Lapped Transform." *IEEE Transactions on Image Processing*, 2, No.1 (1993), 2-17.

[24] Bierling, M. "Displacement Estimation by Hierarchical Blockmatching." *SPIE Visual Commun. Image Proc.*, 1001 (1988), 942-45.

[25] Stiller, C "Motion Estimation for Coding of Moving Video at 8 kbits/s with Gibbs Modeled Vectorfield Smoothing." *SPIE Visual Commun. Image Proc.*, (1990), 468-76.

[26] Ersoy,O. "On Relating Fourier, Sine and Symmetric Cosine Transforms." *IEEE Transactions on ASSP*, 2, No.2 (Feb. 1985), 219-22.

[27] Yip, P. and Rao, K.R. "On the Shift Property the DCT's and DST,s"*IEEE Transac-tions on ASSP*, 35, No.3 (March 1987) 404-406.

[28] Yip, P. and Rao, K.R. "A Fast Computational Algorithm for the Discrete Sine Transform." *IEEE Transactions on Commun.*, 28 (Feb.1980),304-307.

[29] Murthy, N.R. and Swamy, M.N. "On the Computation of Running Discrete Cosine and Sine Transforms." *IEEE Transactions on Signal Processing*, 40, No.6 (June 1992), 1430-37.

[30] Cvetkovic, Z. and Popovic, M.V. "New Fast Recursive Algorithms for the Compu-tation of Discrete Cosine and Sine Transforms." *IEEE Transactions on Signal Processing*, 40, No. 8 (August 1992), 2083-86.

[31] Hou, H.S. "The Fast Hartley Transform Algorithm." *IEEE Trans. on Computers*, C36, No.2 (Feb. 1987) 147-55.

[32] Mahmoud, S.A. Afifi, M.S. and Green, R.J. "Motion Detection and Velocity Comp-utation of Moving Objects of Variable Velocities in Image Sequences." *Proceedings of the Annual Pittsburgh Conference on Modeling and Simulations*, Pittsburgh, (May 5-6,1988), 1095-99.

[33] Nahi, N. and Mora, S.L."Estimation Detection of Object Boundaries in Noisy Images." *IEEE Trans. on Automatic Control*, 13 (1978), 834-45.

[34] Huang, T.S. and Tsai, R.Y."Image Sequence Analysis : Motion Estimation." In:*Image Sequence Analysis* (T.S.Huang, Ed.). New York: Springer-Verlag, 1981.

[35] Mahmoud, S.A. Afifi, M.S. and Green, R.J. "The Effects of Image Background on

Velocity Computation of Moving Objects." *The Journal of Microcomputer Applications*, 13 , No.1 (Jan 1990), 89-96.

[36] Mahmoud, S.A.; Afifi, M.S. and Green, R.J. "Recognition and Velocity Computat-ion of Large Moving Objects in Images." *IEEE Transactions on ASSP*, 36, (Nove-mber 1988), 1790-91.

[37] Rajala, S.A.; Riddle, A.N. and Snyder, W.E."Application of the One Dimensional Fourier Transform for Tracking Moving Objects in Noisy Environments."*Computer Vision, Graphics, and Image Processing*, 21, (1983), 280-93.

تحليل حركة الأجسام باستخدام محول الجيب الرقمي

صبري أ. محمود
قسم هندسة الحاسب الآلي، كلية علوم الحاسب والمعلومات
جامعة الملك سعود ، ص.ب ٥١١٧٨ ، الرياض ١١٥٤٣ ، المملكة العربية السعودية

ملخص البحث . يعرض هذا البحث طريقة جديدة في حساب السرعة وتحليل حركة الأجسام. وتعتمد تلك الطريقة على تطبيق محول الجيب الرقمي على مجموعة من الصور المتعاقبة، كما يعرض هذا البحث المعادلات الرياضية المستخدمة في تطبيق هذا المحول على مجموعة من الصور المتعاقبة والمحتوية على مجموعة من الأجسام المتحركة.

أظهر تحليل طيف هذا المحول للصور المتعاقبة أن سرعة الأجسام المتحركة لها علاقة بمواقع قمم هذا الطيف. وقد تم استخدام خوارزمي لتحديد مواقع قمم طيف المحول الجيبي الرقمي للصور المتعاقبة ومن ثم حساب سرعة الأجسام المتحركة.

تبين نتائج التجارب المعروضة القيمة التطبيقية لهذه الطريقة، فهي سهلة وسريعة وتحسب سرعة الأجسام المتحركة بدقة. ليست هنالك أية فرضيات تضع حداً لحجم أو سرعة الأجسام الممكن قياس سرعتها. وإن وجود الأجسام الثابتة في الصور ليس له أي أثر على هذه الطريقة، كما أن وجود مثل هذه الأجسام لا يتطلب أية معالجة خاصة.